

General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.
- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.
- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.
- This document is paginated as submitted by the original source.
- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

(75-10059) LARGE AREA CROP INVENTORY
EXPERIMENT (LACIE). LANDSAT 3X GAIN STUDY
(NASA) 40 p HC A03/MF A01 CSCL 02C

N79-13466

LACIE-00500

IC-09868

Unclass

G3/43 00059

LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)



75-10059

TM-79924

Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resources Survey
program information and without liability
for any use made thereof.

NASA NOAA USDA

LANDSAT 3X GAIN STUDY



National Aeronautics and Space Administration
LYNDON B. JOHNSON SPACE CENTER

Houston, Texas

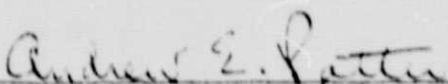
October, 1975



FOREWORD

This document outlines the causes and results of a study of multispectral scanner imagery to compare the advantages of using high-gain (3X) instead of normal-gain (1X) data from the Land Satellites of the Earth Resources Program for the recognition of vegetation for agricultural applications. The Earth Observations Division, Science and Applications Directorate, of the Lyndon B. Johnson Space Center, National Aeronautics and Space Administration, directed the study. Dr. Stanton Yao of Lockheed Electronics Company, Inc., Aerospace Systems Division, Houston, Texas, coordinated the effort under Contract NAS 9-12200 and documented the analysis and results in this report.

APPROVED BY


Andrew E. Potter, Chief
Research, Test, and Evaluation Branch

Original photography may be purchased from:
EROS Data Center

Sioux Falls, SD 57198

CONTENTS

Section		Page
1.0	<u>INTRODUCTION</u>	1
	1.1 RATIONALE BEHIND THE STUDY.	1
	1.2 BACKGROUND.	2
2.0	<u>OBJECTIVES</u>	4
3.0	<u>APPROACH</u>	5
4.0	<u>ANALYSIS AND INTERPRETATION OF RESULTS</u> . . .	14
5.0	<u>CONCLUSION</u>	22
6.0	<u>REFERENCES</u>	24
Appendix		
A	MATHEMATICAL BASIS FOR REDUCING 3X DATA TO 1X AND 2X.	A-1
B	SAMPLE PROGRAM DECK LAYOUT FOR THE GAIN REDUCTION PROGRAM.	B-1
C	PRELIMINARY ANALYSIS OF REGISTRATION ERROR UPON CLASSIFICATION.	C-1

FIGURES

Figure		Page
1	Mean reflectance of winter and spring wheat fields in Hill County (MSS band 4)	6
2	Mean reflectance of winter and spring wheat fields in Hill County (MSS band 5)	7
3	Mean class reflectance for various classes of vegetation in Hill County (MSS band 4) . .	8
4	Mean class reflectance for various classes of vegetation in Hill County (MSS band 5) . .	9
5	Mean reflectance of wheat at the four biological windows for LACIE ITS in Rice County.	10
6	Mean reflectance of west Texas ITS as a function of time.	11
7	High-gain data (MSS band 5, frame 1880-17364)	17
8	Simulated normal-gain data (MSS band 5, frame 1880-17364)	18
9	Class histograms for lettuce at Imperial Valley (MSS bands 4 and 5, frame 1879-17370)	20
10	Clustering of uncalibrated data showing striping effect	21

ABBREVIATIONS AND ACRONYMS

CCT	computer-compatible tape
EOD	Earth Observations Division
ERIM	Environmental Research Institute of Michigan
ERIPS	Earth Resources Interactive Processing System
ERTS	Earth Resources Technology Satellite
GSFC	Goddard Space Flight Center
ITS	intensive test site
JSC	Lyndon B. Johnson Space Center
LACIE	Large Area Crop Inventory Experiment
Landsat	one of two Land Satellites launched by the National Aeronautics and Space Administration in July 1972 and February 1975, respectively (formerly called Earth Resources Technology Satellites)
LARS	Laboratory for Applications of Remote Sensing of Purdue University
LARSYS	classification system developed by the LARS
MLEST	Maximum Likelihood Estimation of Signature Transformation
MSS	multispectral scanner
NASA	National Aeronautics and Space Administration
S&AD	Science and Applications Directorate
SSDA	Sequential Similarity Detection Algorithm

1.0 INTRODUCTION

Prior to the launch of the second Land Satellite (Landsat-2) the Earth Observations Division (EOD), Science and Applications Directorate (S&AD), of the Lyndon B. Johnson Space Center (JSC), National Aeronautics and Space Administration (NASA) undertook a study of multispectral scanner (MSS) data from the first Land Satellite [Landsat-1, formerly called the Earth Resources Technology Satellite (ERTS-1)]. The purpose of the study was to compare the advantages of using the high-gain (3X) data from MSS bands 4 and 5 as opposed to normal-gain (1X) data for agricultural applications such as the Large Area Crop Inventory Experiment (LACIE).

The research involved obtaining ground truth data coincidentally with high-gain MSS data from Landsat-1 covering sites in Imperial Valley, California, on December 19 and 20, 1974. To avoid site dependence, additional high-gain data were gathered over intensive test sites (ITS's) in Kansas on December 27, 28, and 29, 1974. The Landsat-1 MSS data were collected by the Goddard Space Flight Center (GSFC) and shipped to JSC for analysis.

1.1 RATIONALE BEHIND THE STUDY

Most agricultural crops of interest to the LACIE project manifest themselves with reflectances that occupy the lower register of the scales in the visible bands (Landsat MSS bands 4 and 5) during certain times in the growing seasons. Both the sensitivity and the dynamic range of the MSS output would increase using high-gain data with the possible saturation of high-reflectance substances such as snow, clouds,

and bare soil. The increase in sensitivity and in the dynamic range of the data would mean that, for pattern recognition purposes, finer discriminant boundaries for overlapping regions could be defined in measurement space. Thus, it was hypothesized that crop identification accuracies could be improved with high-gain data.

1.2 BACKGROUND

The only analysis of high-gain data known by this author was conducted at the Environmental Research Institute of Michigan (ERIM) (ref. 1). However, this research was not intended to improve classification accuracy but was focused on the saturation characteristics of the high-gain data. The report indicated there were numerous "holes" in the histogram of the two visible bands. This was because the high-gain data was not calibrated at GSFC, the same problem which initially hampered the JSC analysis.

At the time the EOD high-gain study commenced, the goal was to provide some quick results in order to qualify the anticipated sensitivity requirements of Landsat-2 data for LACIE applications. The study was to last only a few weeks; however, no calibrated high-gain data was available until March 31, 1975 (after the launch of Landsat-2). Thus, preliminary conclusions were reached based strictly on the results of the uncalibrated data initially received from GSFC. These results indicated that high-gain data provided no significant classification accuracy improvements over normal-gain data. The negative tone of this conclusion resulted in a downgraded priority for this study, which in turn delayed calibration of the data by GSFC for 3 months.

Fortunately, there was no significant change in the results of the analysis when calibrated data were used.

2.0 OBJECTIVES

The two main objectives of the EOD study of the high-gain data were:

- a. To determine the variations of agricultural crop reflectances as a function of conditions such as growing season, latitude, crop species, and atmospheric conditions, based on an aggregated Landsat-1 MSS analysis previously performed at JSC. In other words, the saturation characteristics of the Landsat-1 data were to be categorized so that, in the event the high-gain data option was exercised for LACIE applications, a strategy as to when and where to use it could be derived.
- b. To demonstrate whether or not, under appropriate conditions and using both the supervised and unsupervised classification approach, the high-gain data would yield improved classification accuracy and proportional estimation for agricultural applications when compared to the normal-gain data.

3.0 APPROACH

To satisfy the first objective of the study, temporally registered Landsat-1 data over Hill County, Montana (high latitude), Rice County, Kansas (medium latitude), and the Texas Panhandle (low latitude) were obtained with and without Sun angle correction. Variations in reflectances of individual crop species from some selected fields throughout the growing season in the Hill County site were also obtained. Figures 1 through 4 are plots of the reflectance counts of MSS bands 4 and 5 as a function of growing season in Hill County; figures 1 and 2 use data from six winter wheat and two spring wheat fields, whereas figures 3 and 4 use averaged data from seven different crops. Figure 5 shows plots of Finney County, Kansas, data, whereas figure 6 shows the Texas Panhandle data, all as a function of growing season.

To satisfy the second objective, the following original experimental design was utilized.

It was determined that the Imperial Valley, an agricultural area usually under clear skies, would be in the overlap region of Landsat-1 MSS images on both December 19 and December 20, 1974. It was anticipated that on December 19 data would be taken under the high-gain option and that the next day normal-gain data would be taken for comparison. A ground truth team was dispatched to the test site on those two dates to take atmospheric transmittance readings and to identify crops and secure crop growth information. However, on December 19, 1974, the data taken over the Imperial Valley

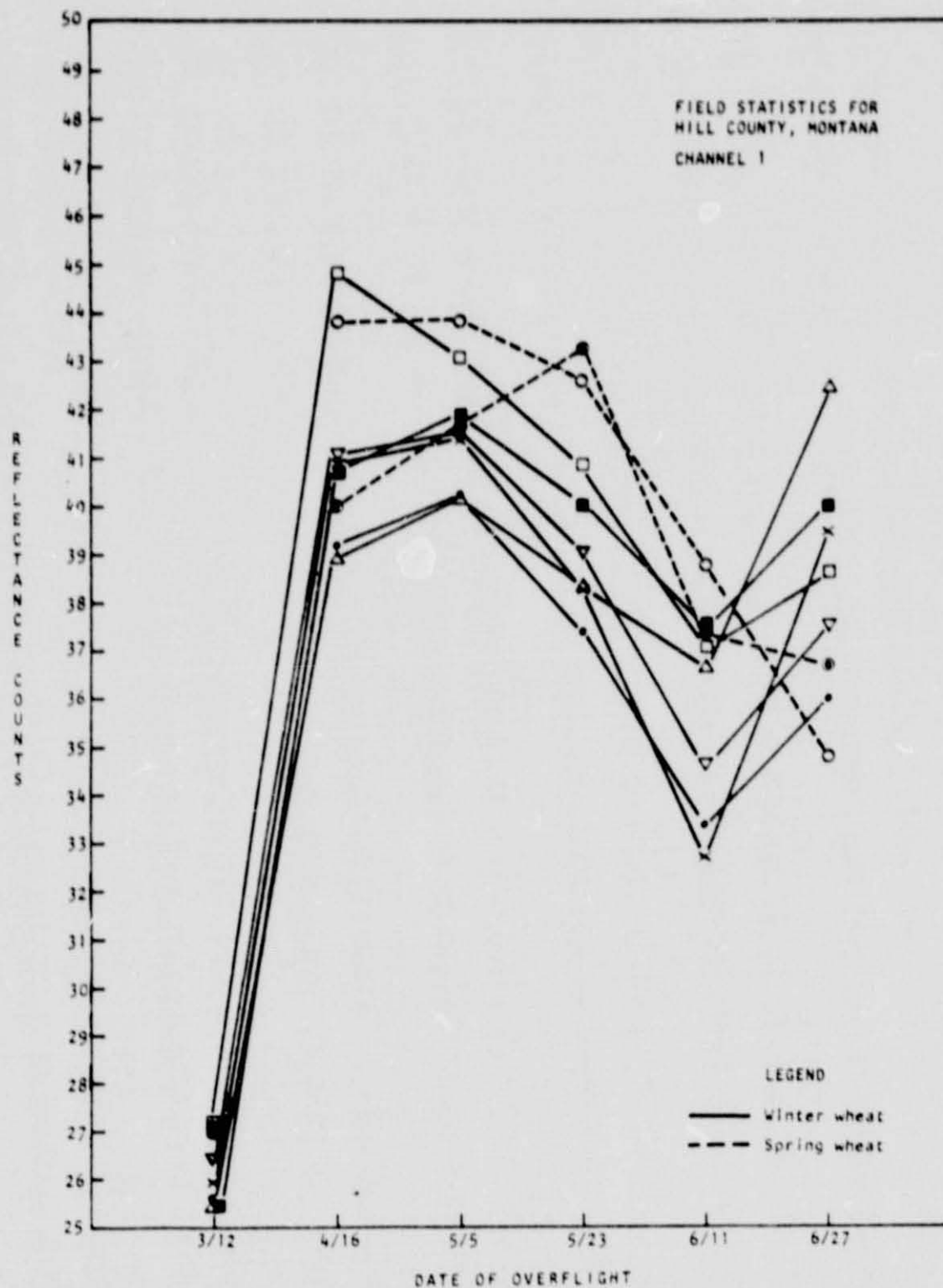


Figure 1.— Mean reflectance of winter and spring wheat fields in Hill County (MSS band 4).

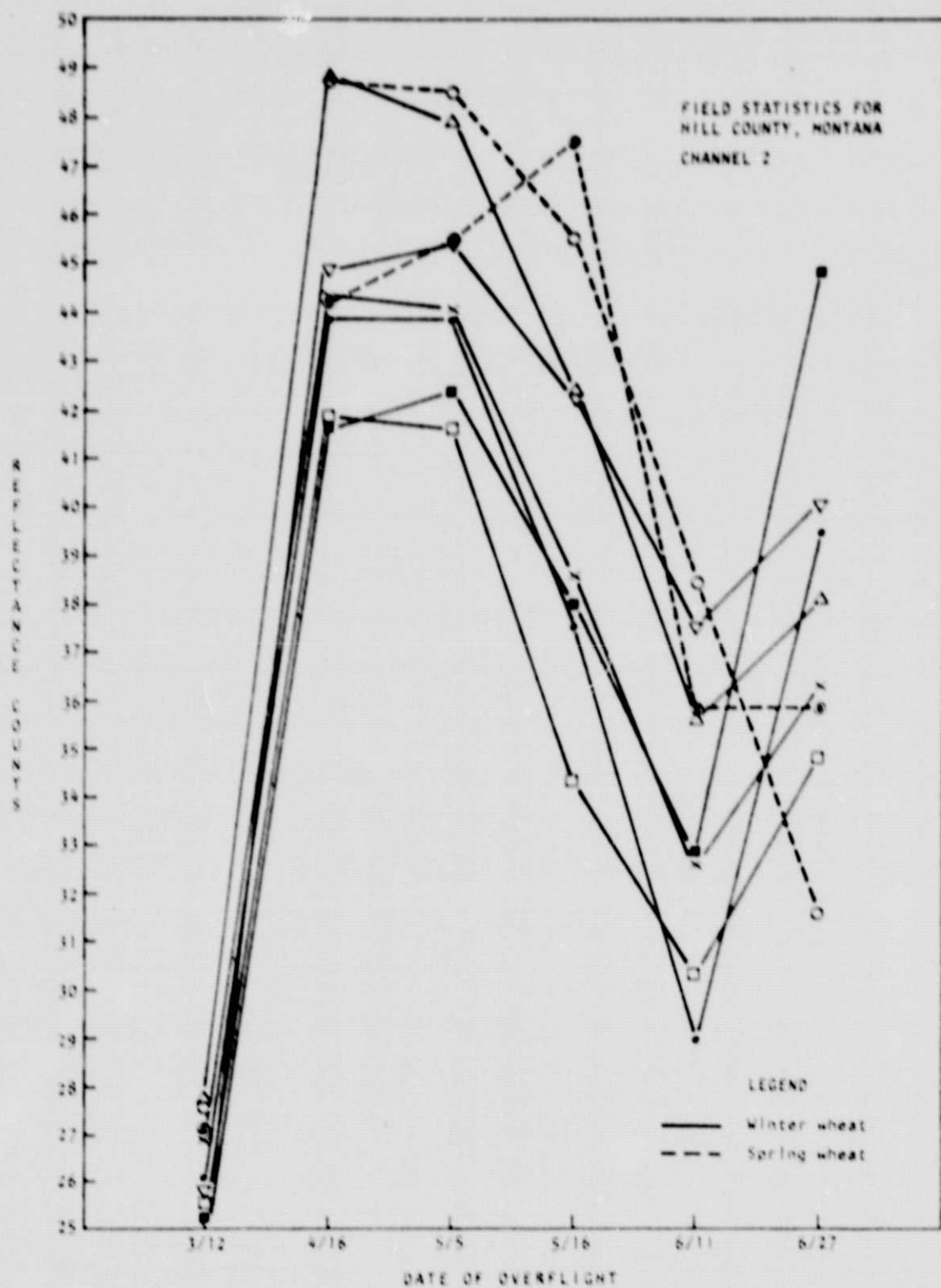


Figure 2.— Mean reflectance of winter and spring wheat fields in Hill County (MSS band 5).

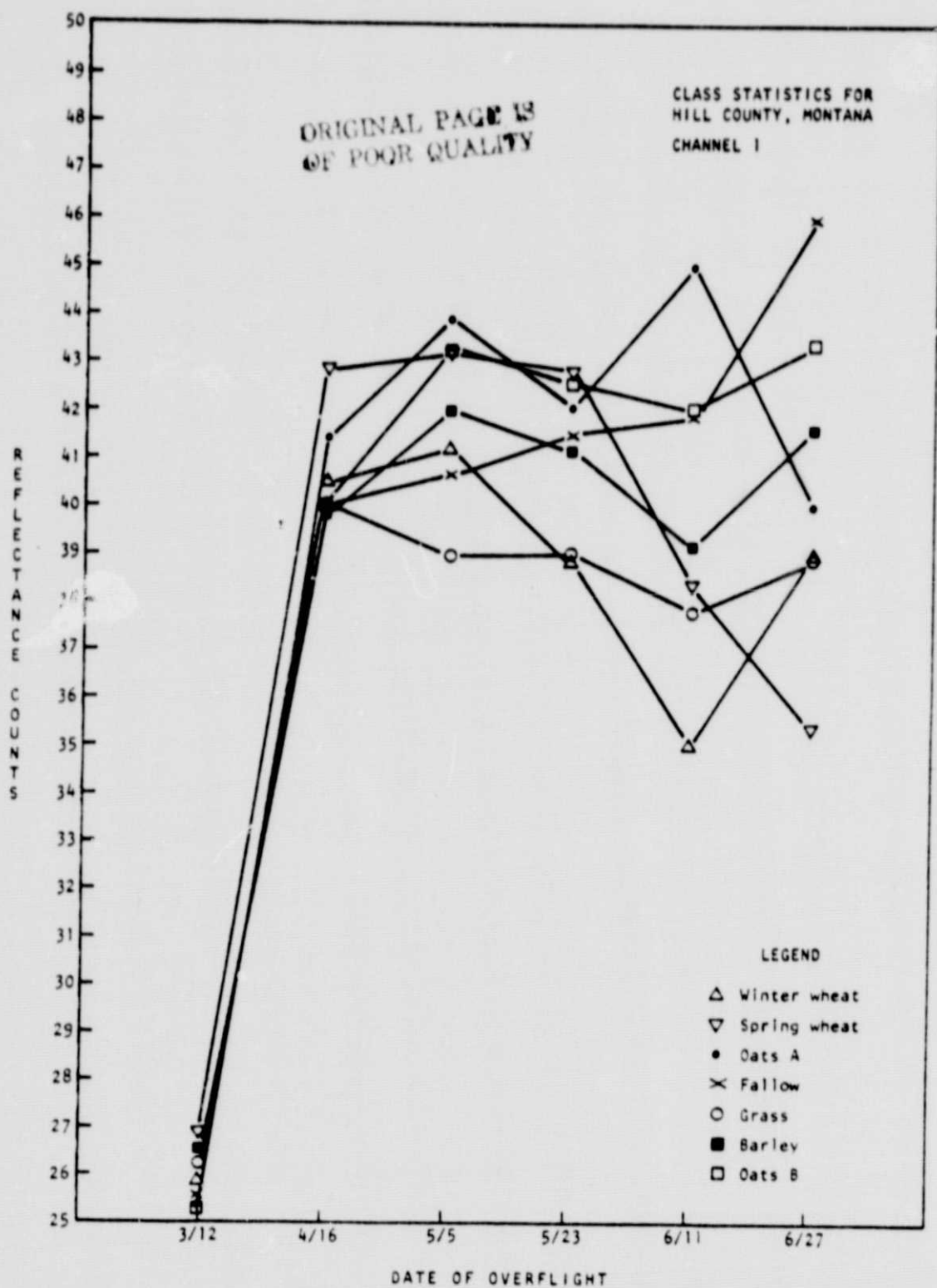


Figure 3.— Mean class reflectance for various classes of vegetation in Hill County (MSS band 4).

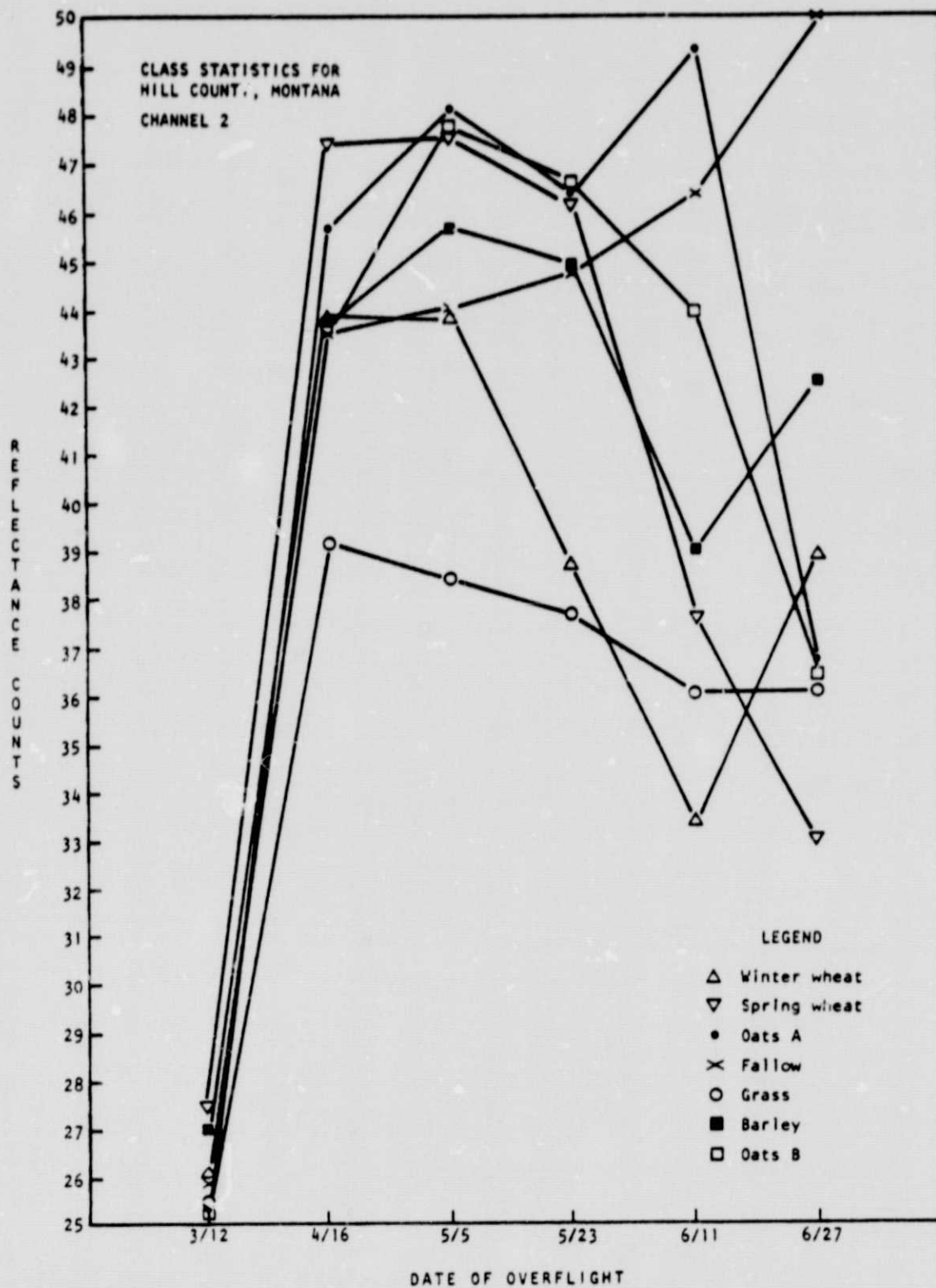


Figure 4.— Mean class reflectance for various classes of vegetation in Hill County (MSS band 5).

ORIGINAL PAGE IS
OF POOR QUALITY

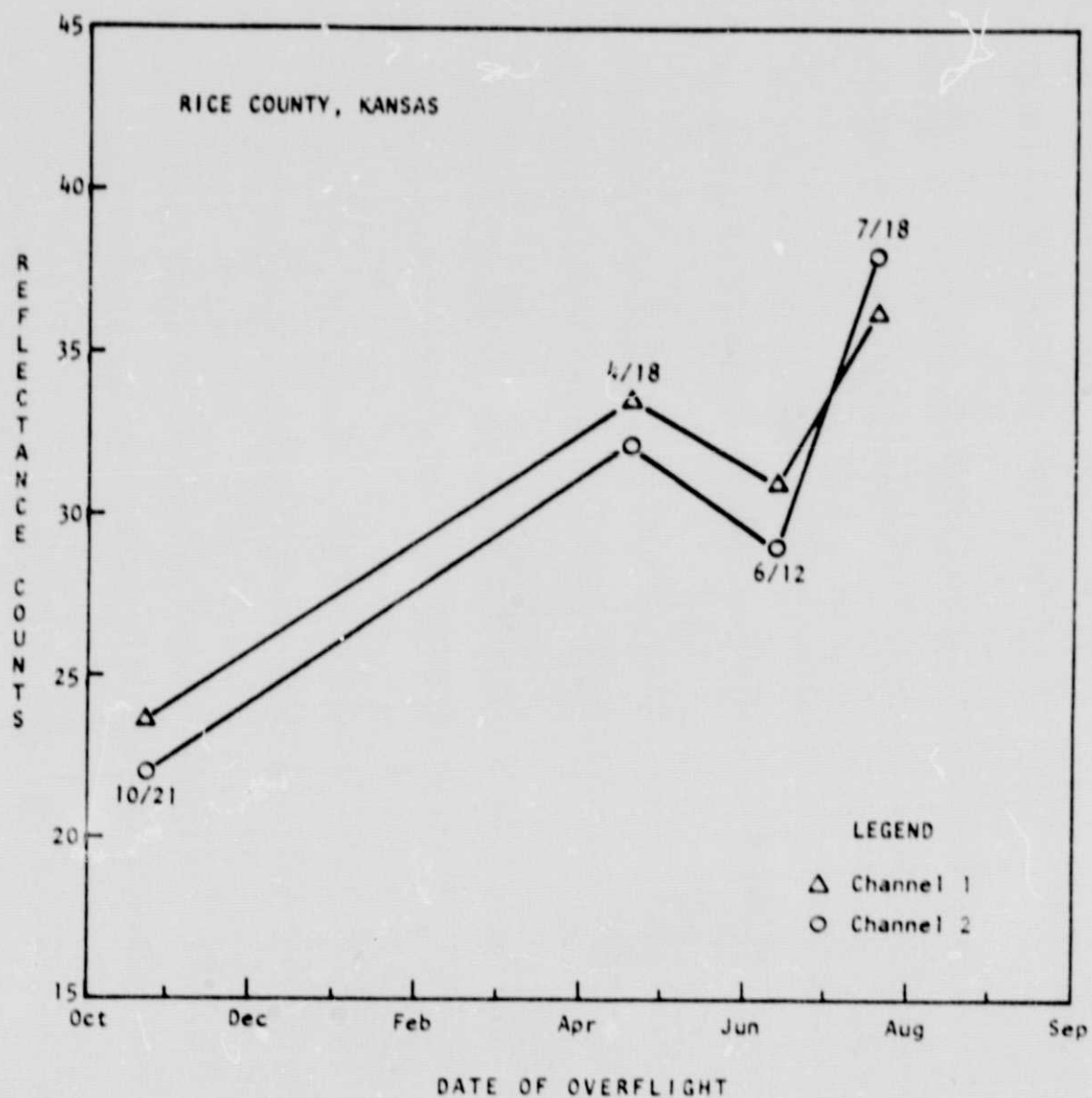


Figure 5.— Mean reflectance of wheat at the four biological windows for LACIE ITS in Rice County.

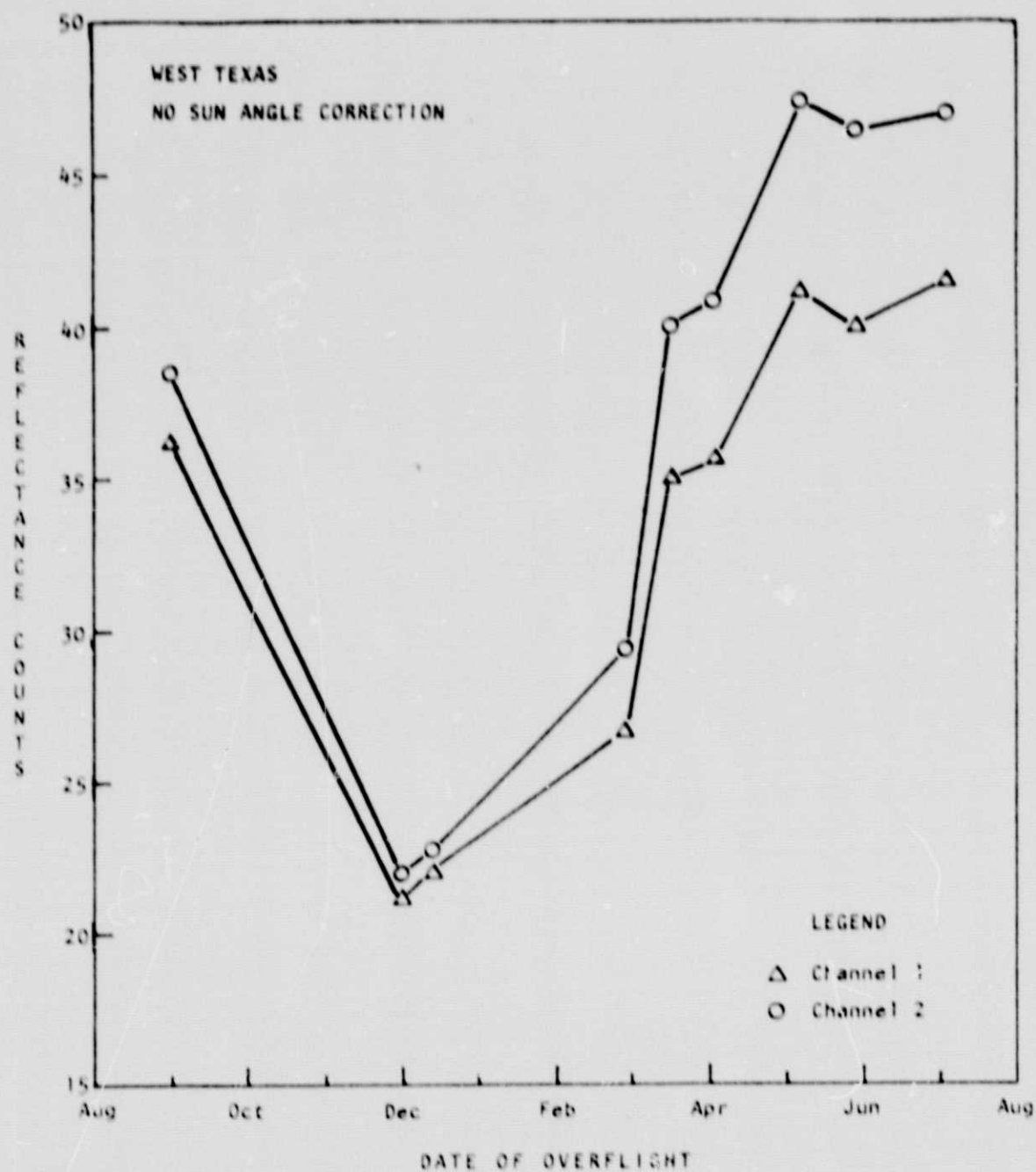


Figure 6.- Mean reflectance of west Texas ITS as a function of time.

was considered unsatisfactory because of smog and haze. The GSFC then proceeded to gather high-gain data on December 20, 1974; as a result, no normal-gain data were collected for comparison. The JSC then devised a method of reducing the high-gain data mathematically to simulate the normal-gain data, as shown in appendix A. Originally, the reduction scheme included a 2X option, but the aforementioned scheme could not simulate 2X data, and this option was excluded from the analysis. The B-matrix multiplication option, which is part of the classification system (LARSYS) of the Laboratory for Applications of Remote Sensing of Purdue University (LARS), was modified to accomplish the reduction technique. (See appendix B for the deck setup used for processing these data.) The high- and normal-gain data tapes from both December 19 and December 20 imagery were then prepared for processing on the Earth Resources Interactive Processing System (ERIPS). The method consisted of the following steps:

- The December 19 high- and normal-gain data were both registered to the December 20 imagery so that a simple set of field definition coordinates could be used on all four image sets (see appendix C).
- Training and test fields were defined and statistics were computed for six classes of interest (wheat, cotton, alfalfa, sugar beets, lettuce, and bare soil). The standard maximum likelihood classifier with assumed equal a priori probabilities was used to classify the test area of interest.
- All four sets of imagery were subjected to a detailed clustering analysis on the ERIPS. An approach which generates class statistics by reading in the even lines

of the image and using them to classify the odd lines of of the image, and vice versa, was planned.

- Color images created from both the high- and normal-gain versions were subjected to the scrutiny of an analyst-interpreter. The addition of human judgment, along with machine processing, for the analysis and comparison of the high- and normal-gain data completed all the goals stated in section 2b.

4.0 ANALYSIS AND INTERPRETATION OF RESULTS

Because of the poor quality of the calibrated data and time allocation problems on the ERIPS, the analysis has not progressed as planned and certain study areas have not been investigated. In addition, as more data were examined, new areas of interest were disclosed. Specific areas of followup analysis are recommended in section 5.

The major result of this investigation of high-gain versus normal-gain Landsat MSS data was not anticipated; that is, the use of high-gain data with its inherent better sensitivity and dynamic range in MSS bands 4 and 5 does not significantly improve Landsat performance for LACIE applications within the context of the stated objectives. The following points support this finding.

- a. The comparison of calibrated and uncalibrated data in table I indicates that any improvement in the classification accuracy of high-gain over simulated normal-gain data is negligible for the six major crop classes considered. The same conclusion was reached when 12 instead of 6 classes were used in the classification or when different a priori probabilities were assigned. To test possible site dependence, one of the Kansas high-gain tapes (Finney County on December 28, 1974) with questionable ground truth data was subjected to the same analysis procedure used for the Imperial Valley data. The identical conclusion was reached. Therefore, insofar as the techniques of supervised pattern recognition are concerned, the high-gain data appears to offer no advantage. It is noteworthy that the statistics in table I indicate that calibration of the December 19 hazy data

TABLE I.- PERCENTAGE OF CLASSIFICATION ACCURACY COMPARISON FOR IMPERIAL VALLEY

Class	December 19, frame 1879-17370 (hazy)				December 20, frame 1880-17364 (clear)			
	Calibrated		Uncalibrated		Calibrated		Uncalibrated	
	3X	1X	3X	1X	3X	1X	3X	1X
Wheat	83.9	83.0	75.2	77.1	83.0	83.9	87.6	87.6
Cotton	85.8	85.2	71.7	71.5	92.0	92.2	92.9	91.7
Alfalfa	41.6	43.8	21.4	17.3	46.6	46.2	46.1	45.0
Sugar beets	63.3	63.3	56.0	57.5	72.2	71.8	64.1	62.9
Lettuce	69.7	68.4	48.7	48.7	77.0	75.9	67.1	70.4
Bare soil	92.9	92.4	85.0	84.5	97.4	97.4	92.2	92.0

greatly improved classification accuracy whereas calibration had little effect on the clear-day data of December 20. However, the relationship between the high- and the normal-gain data remained consistent. A questionable point is the possible results that could be obtained if true instead of simulated normal-gain data were used. Since the two cannot be obtained simultaneously, the effects caused by gain changes and contributing temporal factors must be taken into consideration, which tends to make the analysis more difficult and complicated. Using existing technology, simulation seems to be the best approach.

- b. The analyst-interpreters examined color film copies made from both the high- and the normal-gain Imperial Valley imagery which was taken on December 19 (hazy) and December 20 (clear). They could detect no significant differences in the quality of the imagery. This supports the conclusion that the high-gain imagery is not superior to normal-gain imagery. Figures 7 and 8 are examples of the gray-level images of both the high- and the normal-gain data.
- c. In addition to its lack of improved quality, the high-gain data can be used only sparingly during the winter season with low Sun elevations and at high latitudes; otherwise agricultural targets will be excessively saturated (figs. 1 through 4). Since LACIE imagery must often be gathered on hazy days when the average reflectance can increase substantially, this excessive saturation of agricultural sites becomes even more significant. The December 19 Imperial Valley data was taken under hazy conditions. Even though the Sun elevation was a low 26° ,



Figure 7.- High-gain data (MSS band 5, frame 1880-17364).

ORIGINAL PAGE IS
OF POOR QUALITY



Figure 8.- Simulated normal-gain data (MSS band 5,
frame 1880-17364).

the histogram in figure 9 indicates near saturation (127 counts) for the class of lettuce depicted by MSS band 5.

- d. The class divergence analysis on the Imperial Valley data also showed that MSS band 5 (high gain) and band 7 (normal gain) are the most important bands for the classes considered. Perhaps research using the unsupervised pattern recognition technique of clustering will be able to identify some improvement using high-gain data. Clustering failed when uncalibrated data were used, and the cluster maps showed only excessive striping effects (fig. 10).

HISTOGRAMS FOR FIELDS AND/OR CLASSES
 LETUCE : 1500
 LETUCE : 1520

PP 01 HISTOGRAM.0000

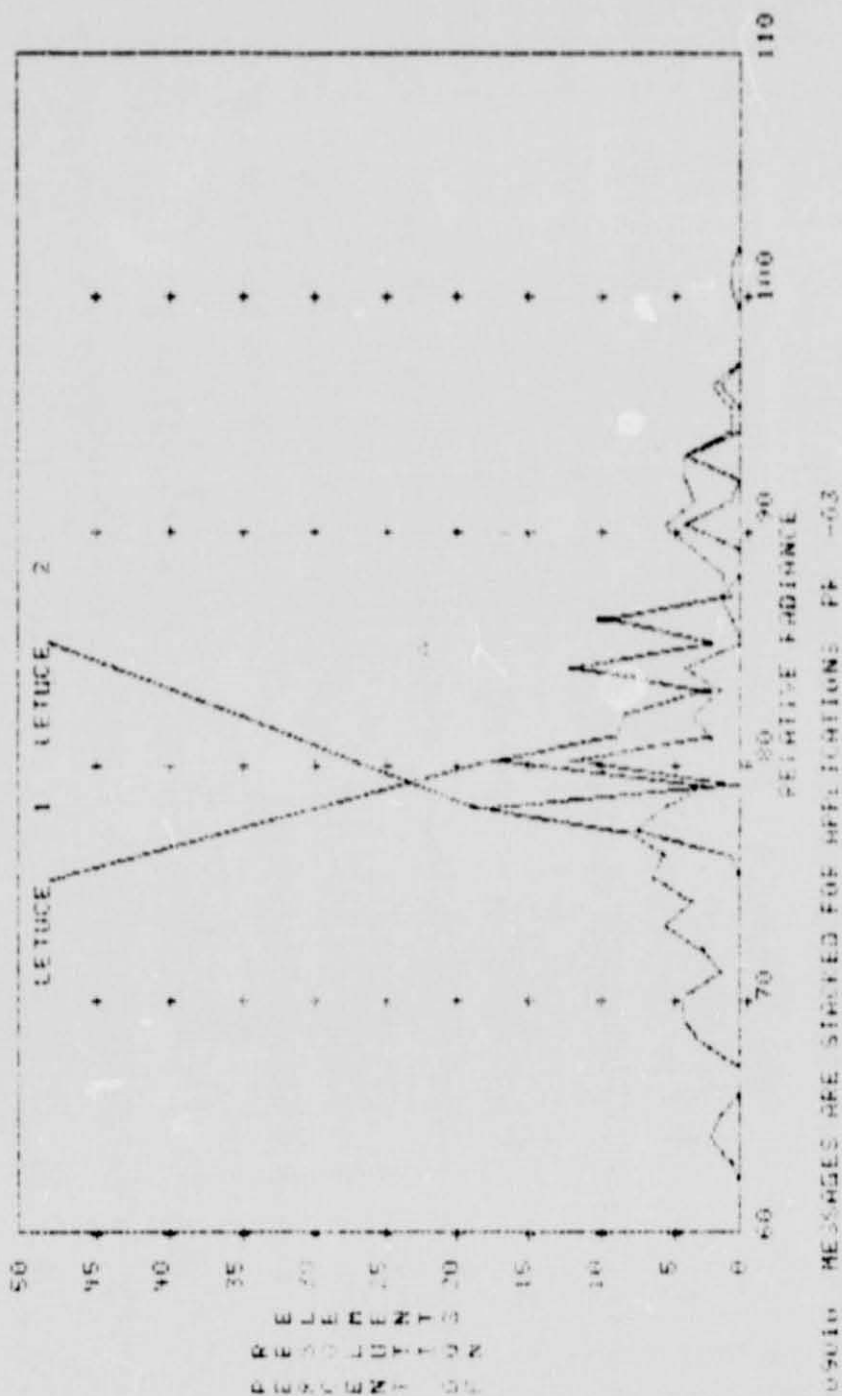


Figure 9.- Class histograms for lettuce at Imperial Valley (MSS bands 4 and 5, frame 1879-17370).

ORIGINAL PAGE 18
OF POOR QUALITY



Figure 10.- Clustering of uncalibrated data
showing striping effect.

5.0 CONCLUSION

In summary, the maximum likelihood classifier on the ERIPS failed to show any improvement in accuracy when comparing the high-gain Landsat data with the simulated normal-gain data. Even if an improvement in accuracy had been detected, the timespan within the crop growing season when the use of high-gain data could be advantageous is limited. It would seem that the high-gain data with their better sensitivities and dynamic ranges would offer some manner of improvement in classification accuracy. However, no improvement with their use has been detected as a result of this analysis.

Because of the lack of time, a thorough study of the recently received calibrated data has not been undertaken. Such an analysis would require new procedures to identify certain characteristics of the clusters which are apparent only in high-gain data and which would indicate that the use of such data might enhance recognition accuracy.

A total of six sets of Landsat-1 imagery are now available for the analysis of gain effects. Three of the sets are high-gain data in MSS bands 4 and 5, whereas the other three are simulated normal-gain data. The four sets obtained over Imperial Valley have good supporting ground truth information for training, whereas the two sets taken over Kansas do not. The use of various combinations of the six data sets is recommended in order to uncover the possible advantages of using the data.

- Method 1 - Use clustering techniques or other appropriate methods to determine whether or not substantial information

that is unavailable with normal-gain data is inherent in high-gain data. If the conclusion is affirmative, a study should be made to determine how this additional information can best be made available for LACIE applications.

- Method 2 - Study the impact of high-gain data on classifiers other than the maximum likelihood classifier, such as the single-class and the two-class classifiers currently being evaluated for LACIE (ref. 2).
- Method 3 - Make use of the existing data sets and the statistical results that have been obtained in the current analysis (for example, the histograms and the field and class means and variances) in order to extend the study into the problem of homogeneity of training field statistics.
- Method 4 - Note that the two sets of Imperial Valley high-gain data, obtained 24 hours apart, were taken under quite different atmospheric conditions. The December 19 images are hazy, whereas the December 20 images are clear. Some readings of the atmospheric transmittances are also available near the test site. The two sets of data could be useful to those interested in the atmospheric effects upon signature extension and indispensable for temporal signature extension to those interested in the various techniques of signature extension. For example, the data sets could be used immediately to test haze correction algorithms such as the Maximum Likelihood Estimation of Signature Transformation (MLEST) techniques (ref. 3).

6.0 REFERENCES

1. Thompson, F.: Analysis of Dynamic Range of ERTS-1 Data Collected with High and Low Gain Options. Report of ERIM (Ann Arbor), 1973.
2. Quirein, J. A.; and Minter, T. C.: Optimal Classification Procedure for the Class M_1 - Not Class M_1 Classification Problem. LEC-1622, Dec. 1973.
3. Levy, S. A.; Minter, T. C.; and Thadani, S. G.: Maximum Likelihood Estimation of Signature Transformation. Technical Memo LEC-4357, May 1975.

APPENDIX A
MATHEMATICAL BASIS FOR REDUCING
3X DATA TO 1X AND 2X DATA

APPENDIX A

MATHEMATICAL BASIS FOR REDUCING 3X DATA TO 1X AND 2X DATA

The available Landsat-1 imagery mentioned in this report was obtained in the high-gain (3X) mode; that is, in MSS bands 4 and 5, the electronic amplification at the sensor output was accelerated three times that in the normal-gain (1X) mode. In this appendix, the mathematical basis for reducing the high-gain data to simulated normal-gain (1X) and double-gain (2X) data is discussed.

Since it is stated in the ERTS handbook that a linear relationship exists between the scene radiance and the data counts obtained from the computer-compatible tape (CCT), the method of reducing the 3X data to 1X is uncomplicated. The analog-to-digital conversion, data compression and decompression, and so forth are not necessary. All that is needed is to divide the data counts by three and truncate:

3X	<u>0,1,2,</u>	<u>3,4,5,</u>	<u>6,7,8</u>	<u>9,10,11</u>
1X	0	1	2	3

3X	...	<u>123,124,125,</u>	<u>126,127</u>
1X		41	42

Thus, the saturation level of 3X data at 127 counts will be reduced to 42 counts in the 1X data, and no data count in 1X data will be greater than 42 counts. However, when reducing 3X data to 2X data, an additional problem arises, as discussed in the unpublished notes on conversion compiled by R. Legault of ERIM.

"Suppose we have 3X gain data with integer counts 1, 2, 3, Simple conversion to 2X gain data involves multiplying by 2/3 and selecting an interval (truncation) rule such as: After multiplication by 2/3, all 3X bins with count strictly less than integer n but greater than or equal to integer count $n - 1$ are named 2X gain-bin count $n - 1$. Multiplication of 3X gain bin counts by 2/3 produces a sequence

$$\begin{array}{ccccccccccc} \underbrace{0, 2/3}, & \underbrace{1-1/3}, & \underbrace{2, 2-2/3}, & \underbrace{3-1/3}, & \underbrace{4, 4-2/3}, & \underbrace{5-1/3}, & \underbrace{6, 6-2/3}, & \underbrace{7}, & \dots \\ 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & \dots \end{array}$$

and use of the above interval (truncation) rule places either one or two 3X bins in a 2X bin. Consequently, a histogram of the 2X simulated data will exhibit the 'missing bin phenomenon' which will impact on classification results.

"The figure below represents the situation interims of analogue signal amplitude.

$$\begin{array}{ccccccc} \text{2X gain bin} & | & n & | & n + 1 & | & n + 2 & | & n + 3 & | \\ & & m & | & m + 1 & | & m + 2 & | & m + 3 & | & m + 4 & | \\ & & & & & & & & & & \text{signal amplitude} \end{array}$$

If the lower signal level of 2X bin n coincides with the lower signal level of 3X bin m (this should be true for $m = n = 0$). Then 3X bin $m + 1$ lies half in 2X bin n and half in 2X bin $n + 1$. Assuming that the signal amplitudes in the 3X bins are uniformly distributed, then the rule for creating 2X bins counts from 3X bins should be: an observation in 3X bins m and $m + 2$ are assigned to 2X bins n and $n + 1$, respectively. An observation in 3X bin $m + 1$ is assigned to 2X bin n with probability 1/2 otherwise bin $n + 1$. Another less satisfactory rule would be to assign the first $m + 1$ observation to 2X bin n , the second $m + 1$ observation to 2X bin $n + 1$ and so on, an alternating rule which puts half of the 3X gain bin $m + 1$ observations into 2X gain bin n and half into 2X gain bin $n + 1$. Either method would take some computing time to do on ERTS frame."

APPENDIX B

SAMPLE DECK LAYOUT FOR THE
GAIN REDUCTION PROGRAM

(Back of deck)

10 FORMAT (4F10.4)

READ(5,10) (CON(1),1=1,4)

READ(5,10) (MIN(1),1=1,4)

READ(5,10) (MAX(1),1=1,4)

-10,11

7 FOR,* DATATR,DATATR

TRI Z

IN Z

TRW Z

ERS

7 XQT CUR

7 ASG K HISFIL

7ANR ASG C=18791

7 ASG Z=Y15815

7S ASG L=SAVE

7N MSG FILE REQ. TAPE 3 FH432 0 FSTRN 0

(Front of deck)

(Back of deck)

SE PMD

\$EXIT

\$END#

1300 1800 1 300 810 1

1.0 1.0 1.0 1.0

0.0 0.0 0.0 0.0

127.0 127.0 127.0 63.0

END

BMATR 1.0

BMATR 1.0 .0 .0 .0 .0

BMATR .33333 .0 .0 .0 .0

BMATR .33333 .0 .0 .0 .0

LCOMB 4NOFET 4YEC 1 2 3 4

B-MATR CARDS

FORMAT OUTPUT=UNIV

\$DATA-TR

7 XQT LARSAA

(Front of deck)

APPENDIX C

PRELIMINARY ANALYSIS OF REGISTRATION
ERROR UPON CLASSIFICATION

APPENDIX C
PRELIMINARY ANALYSIS OF REGISTRATION
ERROR UPON CLASSIFICATION

This brief writeup deals with a preliminary study of the effects of scene misregistration upon classification accuracy based on purely empirical means. In particular, the study gives an upper bound as to the resulting minimizations of variations in accuracy that one can expect when extreme care is taken to ensure "good" registration using Landsat data.

C.1 PROJECT DESCRIPTION

Two sets of digital Landsat-1 imagery (frames 1879-17370 and 1880-17364) were obtained 24 hours apart over Imperial Valley. It is understood that since the two images were taken from adjacent orbits "substantial" rotational misalignment exists between the images. Training and test field boundaries were defined on the data from frame 1880 using ERIPS, and the data from frame 1879 were registered repeatedly (four times) onto the frame 1880 data. The registered data sets from frame 1879 were then classified on ERIPS using the same boundary as defined on the frame 1880 data set. (The resulting accuracy was compared.) The following precautions were taken to ensure that the ERIPS hardware trouble and operation bias would not contribute to the registration error.

- To eliminate operator and screen cursor biases, the Sequential Similarity Detection Algorithm (SSDA) was used for data correlation. It was found that because the two data sets were taken only a day apart the SSDA

worked very well in this situation. The first-order least squares fit based on the SSDA seldom gives more than one-pixel residuals. When a residual of more than one pixel occurs, the correlation point is deleted.

- To ensure that the assignment of a reference data screen would contribute no registration system error, runs 1 and 3 and runs 2 and 4 were assigned different reference screens.
- To totally eliminate cursor positional error, data magnifications from 1 to 3 were used for different registration runs.
- To ensure that correlation points were well distributed over the 500-line by 510-pixel image, as many as 86 points were used for the SSDA and as many as 77 points were entered for least squares computation.

It was anticipated that "perfect" registration would result with all the above precautions taken and that the four classification runs would produce identical results. However, this was not the outcome; the detailed results and some comments are presented in section C.2.

C.2 RESULTS AND CONCLUSIONS

The parameters under which the four registration runs were made are listed in table C-1. The accuracies of the resulting classification on six major classes of crops and soils using identical a priori probabilities and zero threshold values are presented in table C-2. Note that only the relative accuracy among different runs is meaningful here.

TABLE C-1.- REGISTRATION RUN PARAMETERS

Parameter	Run			
	1	2	3	4
Reference screen	2	3	2	3
Point magnification	2	2	3	1
Order of correction polynomial	1	1	1	1
Total correlation points entered	59	62	77	35

TABLE C-2.- CLASSIFICATION ACCURACY COMPARISON

Class	Δ^a	Run			
		1	2	3	4
Wheat	1.4	83.5	82.6	82.1	83.5
Cotton	1.4	81.6	81.8	81.0	82.4
Alfalfa	1.6	51.2	52.2	51.2	52.8
Sugar beets	.8	50.2	50.6	51.0	50.6
Lettuce	.0	59.9	59.9	59.9	59.9
Bare soil	.6	86.8	86.7	87.3	86.7

^aIndicates differences in percentages of classification accuracy between best and worst runs.

The differences in percentage of classification accuracy between the best and the worst runs for all six classes ranged from 0 to 1.6 percent. No particular run shows a clear-cut advantage over all others for all classes, indicating all four registration runs were good but none was outstanding. Thus, it can be concluded that:

- When proper precautions have been taken to do image registration using Landsat data, a difference of about 1 percentage point in classification accuracy cannot be used to indicate the degree of accuracy of the registration.
- For a registered image size of approximately 500 by 500 pixels, it makes no difference for the first-order error approximation whether 35 or 77 correlation points were entered for least squares computation.

Two additional comments are made.

- The differences in percentage of classification accuracy resulting from the four registration runs can be traced to the assignment of certain field boundary points to different classes. Therefore, the possibility existed that by assigning class thresholds to be certain values other than zero, the accuracies of the four runs might be adjusted to be more in line with each other. This method was tried and did not prove to be the case.
- In connection with the adjustments discussed above, the accuracies of the four registration runs might be brought closer together if some interpolation techniques other than the nearest neighbor rule were used during registration. Because of software limitations, it is not possible to investigate this possibility on the ERIPS at the present time.